

Paper 055-2007

Creating Microsoft Excel Pivot Tables from SAS® - (Part 2)

John DeMeo, Royal & SunAlliance, Charlotte NC

ABSTRACT

This paper is a follow-up to a paper presented at SUGI 31 in San Francisco by Ted Conway. It provides several enhancements to a technique that uses Base SAS and Visual Basic Scripting to automate the creation of Excel Pivot Tables from SAS data sets.

INTRODUCTION

The paper presented at SUGI 31 consisted of a macro that produced a delimited file from a SAS data set and then created several pivot tables from the delimited file. The macro worked fine and produced the desired results. I rewrote the macro and included several enhancements such as:

- Provide a list of variables (separated by spaces) to build the delimited file or specify the keyword “_ALL_” to include all the variables in the SAS data set.
- Option to either use the variable names or variable labels. If you specify this option and the label for a variable is blank it defaults to the variable name.
- Specify a file name for the delimited file created.
- Specify a file name for the Visual Basic Script file created.
- Specify the SAS data set for input.
- Character variables retain their leading zeros when opened in Excel.
- Column headings are Bold, Auto Sized and Frozen on the Raw Data Tab.

Note: When the macro builds the Raw Data tab (sheet) it will change all single quotes, double quotes, commas and vertical bars “|” to a single space for the variable labels and data in the spreadsheet.

PIVOT MACRO – SAMPLE CALL

There are several levels to the pivot macro and must be executed in the following order.

Workbook Level: This level builds the delimited file, creates the RAWDATA sheet and starts to create the Visual basic script file.

Specify the input SAS data set

DSN=IN.PIF0605

Specify the name and location of the delimited file created

OUT=D:\JDTEMP\XLWORKBOOK.TXT

Specify the name and location of the Visual Basic Script file created

VB=D:\JDTEMP\XLVB1.VBS

Specify the list of variables for the delimited file or specify the keyword “_ALL_”

VARLIST= “_ALL_”, bring in ALL the variables

OR

VARLIST=STATE BRANCH PREMIUM, use the following list of variable names separated by a space

Specify the level. For this level use WORKBOOK

XLTYPE=WORKBOOK

Specify if you want to use variable labels (Y) or variable names

LBL= Y

Worksheet Level: This level builds the sheet in excel that will contain the pivot table.

Specify the name of the tab (sheet) in excel.

XLNAME=PIVOT1

Field Level: This level builds the guts of the pivot table.

Specify the field name and the field's orientation (Row, Column, Page or Data). You have an optional excel number format and for any data fields an optional statistical measure (Sum, Count Etc).

Resize Level: This level is optional if you wish to auto fit the pivot table's columns.

Create Level: This level creates the workbook.

Here is a sample call. We create the Raw Data tab (sheet) and two pivot tables.

```
%PIVOT(DSN=IN.PIF0605,
        OUT=D:\JDTEMP\XLWORKBOOK.TXT,
        VB=D:\JDTEMP\XLVB1.VBS,
        VARLIST=POLNUM POLSYM INFPREM RATEERR POLTYPE STATE COMPANY
        EXPMONTH PIF,
        XLTYPE=WORKBOOK,
        LBL=Y);

%PIVOT(XLTYPE=WORKSHEET, XLNAME=PIVOT1);
%PIVOT(XLTYPE=FIELD, XLNAME=STATE, ORIENTATION=PAGE);
%PIVOT(XLTYPE=FIELD, XLNAME=COMPANY, ORIENTATION=COLUMN);
%PIVOT(XLTYPE=FIELD, XLNAME=POLTYPE, ORIENTATION=ROW);
%PIVOT(XLTYPE=FIELD, XLNAME=INFPREM, ORIENTATION=DATA, STAT=SUM,
        FORMAT="#,##0.00");

%PIVOT(XLTYPE=RESIZE);

%PIVOT(XLTYPE=WORKSHEET, XLNAME=PIVOT2);
%PIVOT(XLTYPE=FIELD, XLNAME=STATE, ORIENTATION=PAGE);
%PIVOT(XLTYPE=FIELD, XLNAME=COMPANY, ORIENTATION=COLUMN);
%PIVOT(XLTYPE=FIELD, XLNAME=EXPMONTH, ORIENTATION=ROW);
%PIVOT(XLTYPE=FIELD, XLNAME=PIF, ORIENTATION=DATA, STAT=SUM,
        FORMAT="#,##0.00");

%PIVOT(XLTYPE=RESIZE);

%PIVOT(XLTYPE=CREATE);
```

PIVOT MACRO - CODE

```
%MACRO PIVOT(DSN=, OUT=, VB=, VARLIST=, XLTYPE=, XLNAME=, ORIENTATION=, STAT=,
              FORMAT=, LBL=);

%GLOBAL GLOBLBL VBFILE;

%IF %UPCASE(&XLTYPE) EQ WORKBOOK %THEN %DO;
    %LET GLOBLBL=&LBL;
    %LET VBFILE=&VB;
    %LET VARLIST=%UPCASE(&VARLIST);

    %LOCAL CCCC WWWW;

    PROC CONTENTS DATA=&DSN NOPRINT OUT=_TEMP_(KEEP=NAME TYPE VARNUM
                                                LABEL);

    RUN;

    DATA _TEMP_;
        SET _TEMP_;
        NAME=UPCASE(TRIM(LEFT(NAME)));
        LABEL=UPCASE(TRIM(LEFT(LABEL)));
    RUN;

    %IF "&VARLIST" EQ "_ALL_" %THEN %DO;

        DATA _NULL_;
            SET _TEMP_ END=EOF;
            CALL SYMPUT('VAR' || (LEFT(PUT(_N_, 5))), UPCASE(NAME));
            IF EOF THEN CALL SYMPUT('TOTAL', LEFT(PUT(_N_, 8)));
        RUN;
    %END;

    %IF "&VARLIST" NE "_ALL_" %THEN %DO;
        %LET CCCC=1;
        %LET WWWW=%QSCAN(&VARLIST, &CCCC, %STR( ));
        %LET VAR1=%STR(&WWW);
        %DO %WHILE(&WWW NE);
            %LET CCCC=%EVAL(&CCCC+1);
```

```

%LET WWWW=%QSCAN(&VARLIST,&CCCC,%STR( ));
%LET VAR&CCCC=%STR(&WWWW);
%END;
%LET TOTAL=%EVAL(&CCCC-1);
%END;

DATA MV(KEEP=NAME);
  LENGTH NAME $ 32;
  %DO I=1 %TO &TOTAL;
    NAME="&&VAR&I"; OUTPUT;
  %END;
RUN;

PROC SORT DATA=MV NODUPKEY;BY NAME;RUN;

PROC SORT DATA=_TEMP_;BY NAME;RUN;

DATA _TEMP_;
  MERGE MV(IN=INA) _TEMP_(IN=INB);
  BY NAME;
  IF INA AND INB;
  IF LABEL="" OR LABEL=" " THEN LABEL=NAME;
  LABEL=TRANSLATE(LABEL,' ','');
  LABEL=TRANSLATE(LABEL,' ','');
  LABEL=TRANSLATE(LABEL,' ','');
  LABEL=TRANSLATE(LABEL,' ','|');
RUN;

DATA _NULL_;
  SET _TEMP_ END=EOF;
  CALL SYMPUT('VAR' || (LEFT(PUT(_N_,5))),UPCASE(NAME));
  CALL SYMPUT('VARN' || (LEFT(PUT(_N_,5))),_N_);
  CALL SYMPUT('LABL' || (LEFT(PUT(_N_,5))),UPCASE(TRIM(LEFT(LABEL))));
  CALL SYMPUT('TYP' || LEFT(PUT(_N_,5)),TYPE);
  IF EOF THEN CALL SYMPUT('TOTAL',LEFT(PUT(_N_,8)));
RUN;
%GLOBAL MASTTOTAL;
%LET MASTTOTAL=&TOTAL;
%DO I=1 %TO &MASTTOTAL;
  %GLOBAL MASTVAR&I MASTVARN&I MASTLABL&I MASTTYP&I;
%END;
%DO I=1 %TO &MASTTOTAL;
  %LET MASTVAR&I=&&VAR&I;
  %LET MASTVARN&I=&&VARN&I;
  %LET MASTLABL&I=&&LABL&I;
  %LET MASTTYP&I=&&TYP&I;
%END;

DATA _NULL_;
  FILE "&OUT" NOPRINT DLM='|' LRECL=1000;
  IF _N_=1 THEN DO;
    DO IIII=1 TO NOBSLAB;
      SET _TEMP_ NOBS=NOBSLAB POINT=IIII;
      %IF &LBL EQ Y %THEN %DO;
        IF IIII=NOBSLAB THEN PUT LABEL;
        ELSE PUT LABEL @;
      %END;
      %ELSE %DO;
        IF IIII=NOBSLAB THEN PUT NAME;
        ELSE PUT NAME @;
      %END;
    END;
  END;
  SET &DSN;
  FORMAT _NUMERIC_ BEST12.;
  %DO VNUM=1 %TO &TOTAL;
    %IF &&TYP&VNUM=2 %THEN %DO;
      &&VAR&VNUM=TRANSLATE(&&VAR&VNUM,' ','');
      &&VAR&VNUM=TRANSLATE(&&VAR&VNUM,' ','');
      &&VAR&VNUM=TRANSLATE(&&VAR&VNUM,' ','');
      &&VAR&VNUM=TRANSLATE(&&VAR&VNUM,' ','|');
    %END;
  %END;

```

```

        %END;
    %END;
    %DO VNUM=1 %TO &TOTAL;
        PUT &&VAR&VNUM @;
    %END;
    PUT ;
RUN;

%LET XLWBNAME=%QSCAN(&OUT,-2,'\,,:. ');

DATA _NULL_;
    FILE "&VBFILE";

    %LET OUT1=%BQUOTE("&OUT%BQUOTE(");
    PUT 'SET XL = CreateObject("Excel.Application)';
    PUT 'XL.Visible=True';
    PUT 'XL.Workbooks.OpenText "&out1" ', 437, 1, 1, -4142, _ ';
    PUT 'False, False, False, False, False, True, "|", _ ' / 'array( _ ';

    %DO I=1 %TO &TOTAL;
        %IF &I NE &TOTAL %THEN PUT "array(&i,&&typ&i), _ " ;
        %ELSE PUT "array(&i,&&typ&i) _ " ;

        ;
    %END;

    PUT '),True';

    %LET XLWBNAME1=%BQUOTE("&XLWBNAME%BQUOTE(");

    PUT 'XL.Sheets(' "&XLWBNAME1" ').Select';
    PUT 'XL.Sheets(' "&XLWBNAME1" ').Name="RAWDATA"';
    PUT 'XL.Rows("1:1").Select';
    PUT 'XL.Selection.Font.Bold = True';
    PUT 'XL.Range("A2").Select';
    PUT 'XL.ActiveWindow.FreezePanels = True';
    PUT 'XL.Cells.Select';
    PUT 'XL.Selection.Columns.AutoFit';

* For the above array statement, the first dimension is the variable number and
  the second dimension ;
* tells if its numeric 1, or character 2;
%END;

%IF %UPCASE(&XLTYPE) EQ WORKSHEET %THEN %DO;
    PUT "XL.Sheets.Add.name = "&XLname"" /
        "XL.ActiveSheet.PivotTableWizard SourceType=xlbase,
        xl.sheets("&RAWDATA").UsedRange, "&XLname!R1C1", "pvttbl"";
    %END;

%IF %UPCASE(&XLTYPE) EQ FIELD %THEN %DO;

    %IF %UPCASE(&GLOBLBL) EQ Y %THEN %DO;
        %DO I=1 %TO &MASTTOTAL;
            %IF &XLNAME EQ &&MASTVAR&I %THEN %DO;
                %LET XLNAME=&&MASTLABL&I;
                %LET I=%EVAL(&MASTTOTAL+1);
            %END;
        %END;
    %END;

%IF %UPCASE(&ORIENTATION) EQ DATA %THEN %DO;
    PUT "XL.ActiveSheet.PivotTables("&pvttbl").AddDataField
        XL.ActiveSheet.PivotTables("&pvttbl").PivotFields("&XLname"),
        "&Stat of &XLname",
        %IF %UPCASE(&STAT) EQ SUM %THEN "-4157";
        %IF %UPCASE(&STAT) EQ COUNT %THEN "-4112";
        %IF %UPCASE(&STAT) EQ AVERAGE %THEN "-4106";
    %END;

%ELSE %DO;

```

```

        PUT
"XL.ActiveSheet.PivotTables("pvttbl").PivotFields("&XLname").Orientation = "
        %IF %UPCASE(&ORIENTATION) EQ PAGE %THEN "3";
        %IF %UPCASE(&ORIENTATION) EQ ROW %THEN "1";
        %IF %UPCASE(&ORIENTATION) EQ COLUMN %THEN "2";
    %END;

    %IF &FORMAT^= %THEN
    %IF &STAT^= %THEN
        PUT "XL.ActiveSheet.PivotTables("pvttbl").PivotFields
            ("&stat of &XLname").numberformat = " "&format";
    %ELSE
    PUT
"XL.ActiveSheet.PivotTables("pvttbl").PivotFields("&XLname").numberformat = "
"&format";;
    %END;

%IF %UPCASE(&XLTYPE) EQ RESIZE %THEN
    PUT "XL.ActiveSheet.Columns.AutoFit";;

%IF %UPCASE(&XLTYPE) EQ CREATE %THEN %DO;
    RUN;
    X "&VBFILE";
%END;

%MEND PIVOT;

```

PIVOT MACRO – GENERATED VISUAL BASIC SCRIPT FILE

```

SET XL = CreateObject("Excel.Application")
XL.Visible=True
XL.Workbooks.OpenText "D:\JTEMP\XLWORKBOOK.TXT", 437, 1, 1, -4142, _
False, False, False, False, False, True, "|", _
array( _
array(1, 2), _
array(2, 2), _
array(3, 1), _
array(4, 1), _
array(5, 2), _
array(6, 2), _
array(7, 2), _
array(8, 2), _
array(9, 2) _
),True
XL.Sheets("XLWORKBOOK").Select
XL.Sheets("XLWORKBOOK").Name="RAWDATA"
XL.Rows("1:1").Select
XL.Selection.Font.Bold = True
XL.Range("A2").Select
XL.ActiveWindow.FreezePanels = True
XL.Cells.Select
XL.Selection.Columns.AutoFit
XL.Sheets.Add.name = "PIVOT1"
XL.ActiveSheet.PivotTableWizard SourceType=xlbase, xl.sheets("RAWDATA").UsedRange,
"PIVOT1!R1C1", "pvttbl"
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("RISK STATE").Orientation = 3
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("COMPANY").Orientation = 2
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("POLTYPE").Orientation = 1
XL.ActiveSheet.PivotTables("pvttbl").AddDataField
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("INFPREM"), "SUM of
INFPREM",-4157
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("SUM of INFPREM").numberformat =
"#,##0.00"
XL.ActiveSheet.Columns.AutoFit
XL.Sheets.Add.name = "PIVOT2"
XL.ActiveSheet.PivotTableWizard SourceType=xlbase, xl.sheets("RAWDATA").UsedRange,
"PIVOT2!R1C1", "pvttbl"
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("RISK STATE").Orientation = 3
XL.ActiveSheet.PivotTables("pvttbl").PivotFields("COMPANY").Orientation = 2

```

```

XL.ActiveSheet.PivotTables("pvtttbl").PivotFields("EXPMONTH").Orientation = 1
XL.ActiveSheet.PivotTables("pvtttbl").AddDataField
XL.ActiveSheet.PivotTables("pvtttbl").PivotFields("PIF"), "SUM of PIF",-4157
XL.ActiveSheet.PivotTables("pvtttbl").PivotFields("SUM of PIF").numberformat =
"#,##0.00"
XL.ActiveSheet.Columns.AutoFit

```

PIVOT MACRO – OUTPUT

	A	B	C	D	E	F	G	H	I	J	K	L
1	COMPANY	EXPMONTH	INFPREM	PIF	POLICY NUMBER	POLICY SYMBOL	POLTYPE	RATING TERRITORY	RISK STATE			
2	G	200609	2149	1	0002EX	KG	Home	031	IL			
3	A	200702	183	1	0004FC	DU	Auto	99	CA			
4	P	200701	243	1	000580	OB	Other		FL			
5	P	200701	400	1	000607	OB	Other		FL			
6	P	200607	771	1	001495	OB	Other		FL			
7	R	200705	538	1	0021CC	KG	Home	043	NY			
8	R	200705	846	1	0029CG	KG	Home	007	NY			
9	G	200704	1048	1	0031EA	KG	Home	005	NY			
10	D	200703	441	1	0032DH	KG	Home	043	NY			
11	A	200606	865	1	0034EZ	DA	Auto	17	WV			
12	A	200704	764	1	0034FC	DU	Auto	99	CA			
13	G	200702	866	1	0038DK	KG	Home	002	IL			
14	R	200705	959	1	0044BJ	KG	Home	038	FL			
15	A	200607	418	1	0045CL	DU	Auto	115	CT			
16	G	200606	247	1	0046FL	DA	Auto	18	WV			
17	P	200701	334	1	0047DE	KG	Home	087	NY			
18	D	200702	360	1	0047DH	KG	Home	043	NY			
19	R	200610	1228.56	1	0052BY	DA	Auto	32	NY			
20	G	200606	901	1	0052EZ	DA	Auto	18	WV			
21	P	200606	367	1	0065EP	KG	Home	038	IN			
22	P	200607	110	1	006687	LV	Other		MA			
23	P	200606	143	1	007315	LV	Other		NY			
24	P	200607	176	1	007442	LV	Other		MA			
25	A	200703	720	1	0087BX	DU	Auto	99	CA			
26	A	200703	379	1	0089BX	DU	Auto	99	CA			
27	P	200608	795	1	0094CN	KG	Home	030	MD			
28	R	200702	320	1	0098DH	KG	Home	043	NY			
29	P	200610	1688	1	0098ED	DA	Auto	194	CT			
30	P	200606	232	1	0102DO	KG	Home	034	PA			
31	P	200606	528	1	0105EP	KG	Home	037	MD			
32	D	200608	889	1	0107FN	KG	Home	030	RI			
33	P	200609	154	1	011779	LV	Other		MA			
34	A	200610	374	1	0118FW	DU	Auto	36	NY			

Raw Data Tab generated. Note that the leading zeros were kept for Column H. Note we are using the variable labels instead of the variable names where they exist.

RISK STATE	(All)						
Sum of INFPREM	COMPANY						
POLTYPE	A	D	G	P	R	Grand Total	
Auto		535,614.00	250,219.20	303,199.92	504,178.38	630,200.16	2,223,411.66
Home		115,893.00	230,441.00	123,730.00	561,380.00	717,406.00	1,748,850.00
Other					220,555.00	3,500.00	224,055.00
Grand Total		651,507.00	480,660.20	426,929.92	1,286,113.38	1,351,106.16	4,196,316.66

Pivot 1 generated with Risk State as Page Orientation, Poltype as Row Orientation and Company as Column Orientation. Note that Risk State is the variable label for the variable State.

RISK STATE	(All)						
Sum of PIF	COMPANY						
EXPMONTH	A	D	G	P	R	Grand Total	
200606		168.00	72.00	51.00	210.00	177.00	678.00
200607		127.00	60.00	58.00	189.00	139.00	573.00
200608		117.00	48.00	43.00	142.00	151.00	501.00
200609		83.00	40.00	43.00	154.00	143.00	463.00
200610		52.00	28.00	39.00	141.00	132.00	392.00
200611		36.00	36.00	34.00	123.00	122.00	351.00
200612		44.00	24.00	26.00	138.00	130.00	362.00
200701		47.00	22.00	27.00	111.00	118.00	325.00
200702		51.00	31.00	24.00	103.00	97.00	306.00
200703		53.00	21.00	17.00	82.00	46.00	219.00
200704		36.00	18.00	12.00	56.00	63.00	185.00
200705		14.00	23.00	5.00	47.00	45.00	134.00
Grand Total		828.00	423.00	379.00	1,496.00	1,363.00	4,489.00

Pivot 2 generated with Risk State as Page Orientation, Expmonth as Row Orientation and Company as Column Orientation. Note that Risk State is the variable label for the variable State.

CONCLUSION

I want to thank Ted Conway for the original paper. It was his paper that led me to all the enhancements in this paper. Feel free to enhance this macro even more.

REFERENCES

SAS 8.2

Microsoft Excel 2002

Conway, T Sugi 31 Paper 35 – Sur La Table: Creating Microsoft Excel Pivot Tables in a Jiffy from SAS® Data.

CONTACT INFORMATION

Contact the author at:

John DeMeo

Royal & SunAlliance Insurance Company

3600 Arco Corporate Drive

Charlotte, NC 28273

Work Phone: 704.522.3037

Email: J_DeMeo@rsausa.com

TRADEMARKS

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.